

Discrimination among the Human β^A , β^S , and β^C -Globin Genes Using Allele-Specific Oligonucleotide Hybridization Probes

ANNA B. STUDENCKI,¹ BRENDA J. CONNER,² CHAKA C. IMPRAIM,²
RAYMOND L. TEPLITZ,² AND R. BRUCE WALLACE¹

SUMMARY

Synthetic nonadecanucleotides complementary to the human β^A -, β^S -, or β^C -globin sequences were used as hybridization probes to screen human genomic DNA samples for these genes. The oligonucleotides were ³²P-labeled and used as probes to genotype restriction endonuclease digests of human genomic DNA. The data obtained show that hybridization with oligonucleotide probes, unlike restriction fragment length polymorphism (RFLP) analysis or direct restriction enzyme digestion, can be used to directly distinguish among the three alleles of β -globin, β^A , β^S , and β^C , when present either in one (heterozygous) or two copies.

INTRODUCTION

Over the period of the last 5 years, three sophisticated methodologies have been developed to examine human genomic DNAs for genes that give rise to genetic diseases due to point mutations [1–7]. The experimental protocols comprising these methodologies share common elements: each relies on digestion of DNA samples with restriction enzymes, on electrophoretic separation of the fragments that are generated, and on the use of radiolabeled DNAs for probing the resolved digests. However, because each methodology has a unique theoretical basis, important differences among these methodologies are found in the details of protocol execution. These include differences in the strategy of DNA restriction

Received April 5, 1984; revised July 12, 1984.

This work was supported by grant HL29516 from the National Institutes of Health (to R. B. W.). R. B. W. is a member of the Cancer Center at the City of Hope (NIH CA33572).

¹ Department of Molecular Genetics, Beckman Research Institute of the City of Hope, 1450 East Duarte Road, Duarte, CA 91010.

² Division of Cytogenetics and Cytology, City of Hope Medical Center, 1500 East Duarte Road, Duarte, CA 91010.

© 1985 by the American Society of Human Genetics. All rights reserved. 0002–9297/85/3701–0003\$02.00

enzyme digestion and in the size of DNA hybridization probes—long cDNAs or short oligomers—chosen for screening of genomic digests. The sum of these variations determines whether a given methodology approaches the problem of identifying affected genes directly or indirectly. In addition, these variations affect the applicability of a methodology, either limiting or extending its usefulness.

The first of these methodologies is indirect and depends on the association of a given point mutation with a restriction fragment length polymorphism (RFLP). If no such association can be found, this approach is not applicable. Because the frequency of linkage of the point mutation to a particular RFLP is never 100%, this approach requires a family study to determine the pattern of gene segregation among the polymorphic restriction fragments. Unfortunately, unless a second RFLP is found, genotyping of offspring is not possible in families where both the normal and abnormal genes happen to be associated with the same fragment length in the parents [8]. Some point mutations known to be linked with RFLP(s) include: sickle-cell anemia, which is associated with an *HpaI* polymorphism [1, 2, 9] and also with a combined *HpaI*, *HindIII* polymorphism [8, 10]; $\beta^{39(\text{UAG})}$ -thalassemia, which is associated with a *BamHI* polymorphism [11], and β^{C} disease, which is associated with an *HpaI* polymorphism [9].

The second of these methodologies is dependent on the nature of the point mutation itself. If a point mutation creates or destroys a restriction site, genomic DNAs can be examined directly with the enzyme specific for that site. To visualize band patterns in the resolved DNA digests, this methodology, like the first, employs radiolabeled cDNA. The cDNA probe will detect the restriction fragment(s) associated with the normal gene as well as the new restriction fragment(s) associated with the affected gene. Examples of point mutations that are amenable to direct assay in this fashion are: $\beta^{121 \text{ Glu} \rightarrow \text{Lys}}$ (Hb O^{Arab}), which is associated with the loss of an *EcoRI* site [2]; sickle-cell anemia, which is associated with the loss of *DdeI* [4, 5]; and *MstII* [12, 13] sites and $\beta^{\text{IVS-2}}$ -thalassemia, which is associated with the creation of an *HphI* site [14]. This approach is much simpler than the first because it does not require family studies. However, it is also more limited in its applicability.

The newest of the methodologies for detecting point mutations avoids many of the limitations of the first two, as will be illustrated by the work presented here. Unlike the first two methodologies that employ cDNAs as probes, this one employs synthetic oligodeoxyribonucleotides for this purpose. To insure that oligomer probes are of sufficiently high complexity to be specific only for the genes of interest, a minimum length of 19 nucleotide units is chosen for the probes. For each mutation site, a set of nonadecanucleotide probes, which includes at least one oligomer complementary to the normal gene sequence and one complementary to the mutated sequence, is synthesized. A probe complementary to the normal gene will make a duplex with the mutated gene that contains a single base-pair mismatch; the same is true for the duplex formed between the probe for the mutated sequence and the normal gene. Under appropriately stringent hybridization conditions, perfectly matched and mismatched duplexes can be distinguished easily by the presence or absence of detectable signal in the particular

restriction fragment that carries these genes. Examples of affected genes that have been distinguished from their normal counterparts by this approach are the sickle-cell [6, 7], α_1 -antitrypsin [15], $\beta^{39(\text{UAG})}$ -thalassemia [16], and $\beta^{\text{IVS-1}}$ -thalassemia [17] genes.

To illustrate further the power, generality, and simplicity of the oligonucleotide hybridization assay for point mutations, we describe here the use of oligonucleotide probes to distinguish among three alleles of the β -globin gene (β^A , β^S , and β^C). Discrimination of all three of these genes by RFLP analysis is quite laborious and not always possible because of the lack of linkage to an appropriate RFLP. It is impossible by direct restriction analysis. An A \rightarrow T transversion in the sixth codon of the β -globin gene (GAG \rightarrow GTG) produces the $\beta^A\rightarrow\beta^S$ mutation, while a G \rightarrow A transition in this same codon (GAG \rightarrow AAG) produces the $\beta^A\rightarrow\beta^C$ mutation. The β^C mutation does not affect the *DdeI* (CTNAG) or *MstII* (CCTNAGG) restriction enzyme sites and, therefore, unlike the β^S gene, cannot be discriminated from the β^A gene by direct digestion with these enzymes.

MATERIALS AND METHODS

The solid-phase triester method [18] was used to synthesize nonanucleotides, one deca-nucleotide, and nonadecanucleotides (tables 1 and 2). The synthetic nonadecanucleotides 19A, 19S, and 19C are complementary to the coding strands of the β^A -, β^S -, and β^C -globin genes, respectively, while the nonadecanucleotides 19A', 19S', and 19C' are complementary to the noncoding strands of this same set of corresponding genes. Each oligonucleotide makes at least one mismatch with its noncomplementary β -globin genes: 19A with the β^S and β^C sequences, etc.

Probes were either 5' end-labeled with ^{32}P using T4 polynucleotide kinase or internally labeled with ^{32}P using the repair activity of *E. coli* DNA polymerase I (Klenow fragment) and α - ^{32}P deoxyribonucleoside triphosphates as described [6, 7]. Specific activities of the probes are given in the legends of figures 1 and 2.

Human genomic DNAs were prepared from leukocytes obtained from normal subjects, $\beta^A\beta^A$, and from subjects who were heterozygous or homozygous for the sickle-cell and β^C genes (i.e., from subjects with the $\beta^A\beta^S$, $\beta^S\beta^S$, $\beta^C\beta^S$, and $\beta^C\beta^C$ genotypes). DNA from the clone λHBG1 that contains the normal human β -globin gene was prepared as described [19]. A pBR322 plasmid containing the entire normal human β -globin gene was subjected

TABLE 1
SEQUENCES OF OLIGONUCLEOTIDE PRIMERS

Primers	Sequences
9A	5' CTCCTGAGG 3'
9S	CTCCTGTGG
9C	CTCCTAAGG
p10A',S',C'	pGCAGACTTCT*

NOTE: The 9A, 9S, and 9C primers possessed 5' hydroxyl groups as shown or 5' dimethoxytrityl (DMT) protecting groups. The DMT protecting group is useful in probe product purification and does not interfere with hybridization.

* A universal primer that was used to prime synthesis of the internally labeled oligomers p19A', p19S', and p19C' on the templates 19A, 19S, and 19C.

since DNA is double-stranded, two probes can be synthesized for each gene, one specific for each of the two strands (see table 2). The decision as to which of the two possible probes is preferable is dependent on the types of mismatches each probe makes with its two noncomplementary β -globin genes. On theoretical grounds, different base-pair mismatches (e.g., A-A, T-T, G-T, etc.) at a given site in a nonadecamer-genomic DNA duplex might be expected to destabilize that duplex to different degrees. Although the relative effects of all the possible mismatches on duplex stability have not been determined as yet, it was assumed that under the hybridization conditions used, 53°C and 0.9 M Na⁺, the G-T mismatch would have the least effect on the melting temperature of a nonadecamer-genomic DNA duplex; that is, a probe forming a G-T mismatch with a genomic sequence might be expected to produce a signal with that noncomplementary DNA as well as with its complementary sequence. Therefore, in choosing probes to distinguish all three β -globin genes (the β^A , β^S , and β^C DNAs), it was decided to use the 19A', 19C, and 19S or 19S' probes and to eliminate the use of the 19A and 19C' probes that make a G-T mismatch with the β^C and β^A sequences, respectively (see table 2).

The autoradiograms for a single agarose gel that was hybridized with 5' end-labeled nonadecamer probes specific for the β^A , β^S , and β^C genes are shown in figure 1. Ten different human genomic DNA samples were loaded on the gel shown. The samples of the λ H β G1 and pBR322-H β S marker DNAs that were also included on this gel contained 10×10^{-3} fmol of β -globin sequence complementary to the β^A and β^S nonadecamer probes, respectively. This amount is equivalent to that expected in 10 μ g of genomic DNA homozygous for either of these sequences.

The β -globin genotype of each sample as determined by allele-specific hybridization with these probes is given above each sample lane in figure 1. The results of this analysis agreed with those obtained by conventional analysis of blood globin proteins. The correct identification of these genomic DNA samples and marker DNAs shows the specificity of this method for detecting point mutations. Under the conditions chosen for hybridization, the nonadecamer probes hybridized only to their complementary β -globin DNAs, the p19A' probe with β^A DNA, p19S with β^S DNA, and p19C with β^C DNA.

A comparison of the 1.8-kb bands within panel A and also of those within panel B of figure 1 shows that signal intensity for homozygous DNA samples is approximately twice that of heterozygous samples. The > 1.8-kb DNA band seen in each of the lanes 2–11 in panel C is due to an unknown DNA sequence that is complementary to the p19C probe. This signal is unlikely to be due to probe hybridization with any one of the related β -globin sequences (ϵ , γ^G , γ^A , or δ) since the p19C probe makes at least two mismatches with each of these genes (see table 2) and therefore should not hybridize under the conditions used.

Figure 2 shows the autoradiograms obtained for three mini-gels, each of which was loaded with 1- μ g samples of five different human genomic DNAs plus appropriate amounts of λ H β G1 and pBR322-H β S marker DNAs (0.006 and 0.06 pg each of the β^A and β^S marker sequences). Each gel was hybridized with one radiolabeled nonadecamer probe. Since both the 19S and 19S' probes make de-

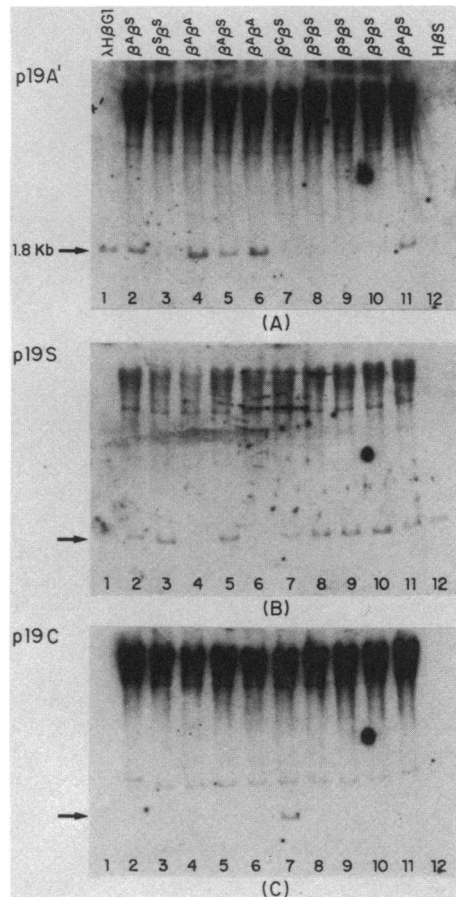


FIG. 1.—Direct gel hybridizations of human genomic DNAs with the 5' ^{32}P end-labeled nonadecamer probes p19A', p19S, and p19C. Ten 10- μg samples of human genomic DNAs that had been digested with *Bam*HI were resolved on a large vertical agarose gel. *Bam*HI digests of 150 pg $\lambda\text{H}\beta\text{G1}$ and of 30 pg pBR322-H βS samples were also included on this gel. Electrophoresis was carried out at 40 V for 16 hrs. This gel was hybridized sequentially with the three probes indicated. Hybridization mixes contained 3.5×10^6 cpm p19A', 2.7×10^6 cpm p19S, or 5.8×10^6 cpm p19C probe/ml mix plus 0.9 M NaCl, 0.18 M Tris-HCl, pH 8.0, 6 mM EDTA, 0.5% Nonidet P-40, 0.1% each of bovine serum albumin, Ficoll, polyvinyl-pyrrolidone and sodium dodecyl sulfate, and 10% dextran sulfate. After each 2-hr hybridization at 53°C, the gel was washed three times for 15 min at 0°C in $6 \times \text{SSC}$ (0.9 M NaCl, 0.09 M sodium citrate, pH 7.2), frozen overnight at -70°C , washed once for 1 min at 53°C in $6 \times \text{SSC}$, and autoradiographed. Autoradiography times were 10, 12, and 4 days for panels A, B, and C, respectively. To remove one probe in preparation for hybridization with the next probe, the gel was washed once for 1 min at 65°C in $6 \times \text{SSC}$. The autoradiography times for the hybridizations shown in panels A and B are long only because the specific activities of the p19A' and p19S probe preparations used were low due to the fact that they had been prepared several days earlier. With probes of specific activity of 2×10^9 cpm/ μg , autoradiography time is typically 2–4 days as in panel C.

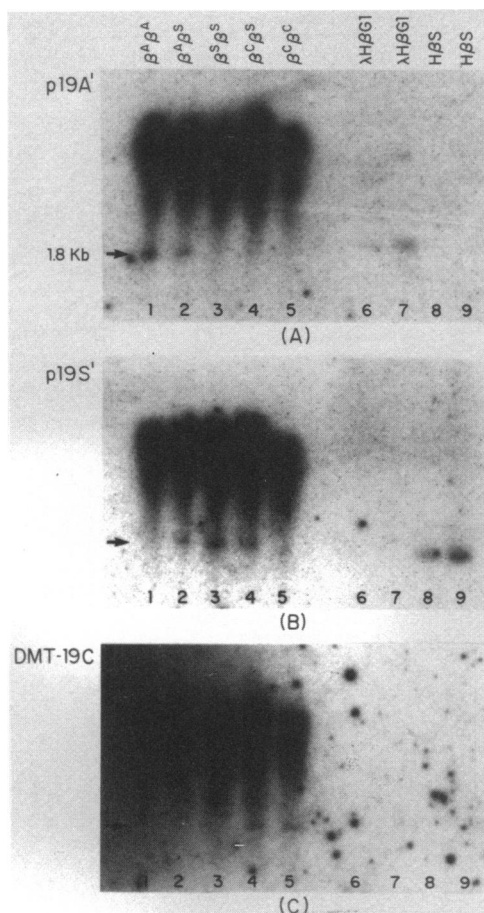


FIG. 2.—Direct gel hybridizations of human genomic DNAs with the ^{32}P primer extended nonadecanucleotide probes p19A', p19S', and DMT-19C. One- μg samples of human genomic DNAs, with the β -globin genotypes given above, were digested with *Bam*HI and *Eco*RI and resolved on each of three vertical 1% agarose mini-gels. *Bam*HI digests of 16 and 160 pg $\lambda\text{H}\beta\text{G1}$ samples and of 3 and 30 pg pBR322-H βS samples were included as markers on each gel. Electrophoresis was done for 2 hrs at 100 V. As indicated above, each gel was hybridized with a different probe. Hybridization mixes contained 3.0×10^6 cpm ^{32}P -labeled probe/ml mix, 10 $\mu\text{g}/\text{ml}$ sonicated, denatured *E. coli* carrier DNA, 0.3% SDS, and $5 \times$ SSPE ($1 \times$ SSPE = 10 mM sodium phosphate, pH 7.0, 0.18 M NaCl, 1 mM EDTA). The specific activities of the p19A', p19S', and DMT-19C probes were 1.48, 1.55, and 1.46×10^{10} dpm/ μg , respectively. The gels in panels A, B, and C were hybridized at 53°C for 23, 24, and 22 hrs. After hybridization, the gels were washed once for 15 min at 0°C in $2 \times$ SSPE and once for 30 min at room temperature in $2 \times$ SSPE, 0.1% SDS. Then the gels in panels A, B, and C were autoradiographed for 2.8, 2.8, and 4.8 days, respectively.

stabilizing mismatch base pairs with their noncomplementary β -globin genes (i.e., neither makes a lone G-T mismatch with any noncomplementary β -globin sequence), the β^S probe chosen for this experiment was the alternate nonadecamer: p19S'. The 10-fold higher specific activities achieved by the internal labeling procedure allowed the analysis of 1 μg of genomic DNA rather than 10 μg of DNA as in figure 1.

As in the experiment with 5' end-labeled probes, the results of genotyping that were obtained with primer extended probes agreed with those of blood globin protein analysis. The correct identification of genomic samples and marker DNAs by direct mini-gel hybridization with these probes of higher specific activity again shows that oligomer probes are allele specific under appropriate hybridization conditions.

In addition, the mini-gel data show that this method of detecting point mutations is very sensitive. When oligomer probes with specific activities of 10^{10} dpm/ μ g are used, gene dosage can, in fact, be determined in 1- μ g samples of genomic DNAs. An examination of the three gels in figure 2 shows that the signal intensity of the 1.8-kb band is approximately two times greater in the lanes with homozygous genomic DNA samples: with the p19A' probe, the $\beta^A\beta^A$ signal is more intense than the $\beta^A\beta^S$ signal; with the p19S' probe, $\beta^S\beta^S > \beta^A\beta^S$ and $\beta^C\beta^S$; and with the DMT-19C probe, $\beta^C\beta^C > \beta^C\beta^S$.

This experiment was designed so that three probes prepared at the same time could be used simultaneously. Obviously, any one of these gels could have been hybridized sequentially with all three probes. Thus, the genotyping could have been accomplished with only 1 μ g rather than 3 μ g of each genomic DNA.

DISCUSSION

We have shown here how allele-specific oligonucleotide probes can be used to distinguish the β^A -, β^S -, and β^C -globin genes unambiguously in restriction digests of total human genomic DNAs. This group of genes was chosen for this study to illustrate the superiority of allele-specific oligonucleotide hybridization over the other two methodologies developed for detecting point mutations. Neither RFLP analysis nor direct restriction analysis is satisfactory for distinguishing among all three of these genes.

The β -globin gene has been associated with an *HpaI* polymorphism that produces 7.0-, 7.6-, or 13.0-kb restriction fragment lengths. However, neither the normal β^A -globin gene nor either of the two variant genes, β^S and β^C , is exclusively associated with any one of these fragment lengths. The β^A gene is most often linked with the 7.6-kb fragment (88% frequency), while both the β^S and β^C genes are usually linked with the 13.0-kb fragment, the β^S gene at a frequency of 68% and the β^C gene at a frequency of $> 95\%$ [5, 8]. Because of this incomplete linkage pattern, there are families for whom it is not possible to establish a pattern of inheritance for these sequences by using the *HpaI* RFLP alone. For example, if the parents have the $\beta^A\beta^S$ and $\beta^C\beta^S$ genotypes and both β^S genes and the β^C gene are linked with 13.0-kb fragments, offspring with a 7.6/13.0-kb restriction pattern may have either the $\beta^A\beta^S$ or $\beta^A\beta^C$ genotype. Those possessing a 13.0/13.0-kb restriction pattern may have the $\beta^C\beta^S$ or $\beta^S\beta^S$ genotype. Similar difficulties can arise in genotyping offspring of parents with $\beta^A\beta^S$ and $\beta^A\beta^C$ genotypes. In such cases, a second analysis of each family member's DNA with a different restriction enzyme that is known to have polymorphic sites in the globin-gene cluster (e.g., *HindIII*) must be done. The details of such extended analyses, using both *HpaI* and *HindIII* polymorphic sites, have been published for simpler cases where the genotypes of individuals who stood to inherit only the β^A and β^S genes

were determined. The results of those experiments showed that the fraction of couples for whom amniocentesis could provide an exact diagnosis of the sickle-cell state was increased from 36% to 80% when the pattern of *Hind*III linkages was also examined [8]. Obviously, twice the DNA is required for such extended analyses and the work involved in genotyping is doubled.

On the other hand, direct analysis of genomic DNAs with restriction enzymes cannot distinguish among individuals with the $\beta^A\beta^A$, $\beta^A\beta^C$, and $\beta^C\beta^C$ genotypes or between individuals with the $\beta^A\beta^S$ and $\beta^C\beta^S$ genotypes because the $\beta^A \rightarrow \beta^C$ mutation does not affect a known restriction site. Furthermore, in rare cases where the $\beta^{\text{Thal(codon 6)}}$ gene is involved, *Mst*II digestion would not allow discrimination of this gene and β^S [21]. Therefore, of RFLP, direct restriction enzyme analysis, and allele-specific hybridization with oligonucleotide probes, the latter is the only way of distinguishing the β^A , β^S , and β^C genes in all cases. Extensive family studies are not necessary because each individual's DNA can be examined and characterized directly.

From the work described here, it can be inferred that allele-specific hybridization with oligonucleotide probes is presently the only methodology that is universally applicable for discriminating among genes that differ by a single base. The necessary oligonucleotide probes can be synthesized for any affected gene once its normal and altered sequences are known.

ACKNOWLEDGMENTS

We thank Marlene Neerman for typing the manuscript.

REFERENCES

1. KAN YW, DOZY AM: Antenatal diagnosis of sickle cell anaemia by DNA analysis of amniotic-fluid cells. *Lancet* ii:910-912, 1978
2. PHILLIPS JA III, SCOTT AF, KAZAZIAN HH JR, STETTEN G, THOMAS GH: Prenatal diagnosis of hemoglobinopathies by restriction endonuclease analysis: pregnancies at risk for sickle cell anemia and S-O^{Ara}b disease. *Johns Hopkins Med J* 145:57-60, 1979
3. KAN YW, DOZY AM: Polymorphism of DNA sequence adjacent to human β -globin structural gene: relationship to sickle mutation. *Proc Natl Acad Sci USA* 75:5631-5635, 1980
4. GEEVER RF, WILSON LB, NALLESETH FS, MILNER PF, BITTNER M, WILSON JT: Direct identification of sickle cell anemia (SS) by blot hybridization. *Proc Natl Acad Sci USA* 78:5081-5085, 1981
5. CHANG JC, KAN YW: Antenatal diagnosis of sickle cell anaemia by direct analysis of the sickle mutation. *Lancet* ii:1127-1129, 1981
6. CONNER BJ, REYES AA, MORIN C, ITAKURA K, TEPLITZ RL, WALLACE RB: Detection of sickle cell β^S -globin allele by hybridization with synthetic oligonucleotides. *Proc Natl Acad Sci USA* 80:278-282, 1983
7. STUDENCKI AB, WALLACE RB: Allele specific hybridization using oligonucleotide probes of very high specific activity: discrimination of the human β^A and β^S -globin genes. *DNA* 3:7-15, 1984
8. PHILLIPS JA III, PANNY SR, KAZAZIAN HH JR, BOEHM CD, SCOTT AF, SMITH KD: Prenatal diagnosis of sickle cell anemia by restriction endonuclease analysis: *Hind* III polymorphisms in γ -globin genes extend test applicability. *Proc Natl Acad Sci USA* 77:2853-2856, 1980
9. KAN YW, DOZY AM: Evolution of the hemoglobin S and C genes in world populations. *Science* 209:388-391, 1980

10. ANTONARAKIS SE, BOEHM CD, GIARDINA PJV, KAZAZIAN HH JR: Non-random association of polymorphic restriction sites in the β -globin gene cluster. *Proc Natl Acad Sci USA* 79:137–141, 1982
11. KAN YW, LEE KY, FURBETTA M, ANGIUS A, CAO A: Polymorphism in DNA sequence in the β -globin region: application to prenatal diagnosis of β^0 thalassemia in Sardinia. *N Engl J Med* 302:185–188, 1980
12. CHANG JC, KAN YW: A sensitive new prenatal test for sickle-cell anemia. *N Engl J Med* 307:30–32, 1982
13. ORKIN SH, LITTLE PFR, KAZAZIAN HH JR, BOEHM CD: Improved detection of the sickle cell mutation by DNA analysis. *N Engl J Med* 307:32–36, 1982
14. BAIRD M, DRISCOLL C, SCHREINER H, ET AL.: A nucleotide change at a splice junction in the human β -globin gene is associated with β^0 thalassemia. *Proc Natl Acad Sci USA* 78:4218–4221, 1981
15. KIDD VJ, WALLACE RB, ITAKURA K, WOO SLC: α_1 -Antitrypsin deficiency detection by direct analysis of the mutation in the gene. *Nature* 304:230–234, 1983
16. PIRASTU M, KAN YW, CAO A, CONNER BJ, TEPLITZ RL, WALLACE RB: Prenatal diagnosis of β thalassemia: direct detection of a single nucleotide mutation in DNA. *N Engl J Med* 309:284–287, 1983
17. ORKIN SH, MARKHAM AF, KAZAZIAN HH JR: Direct detection of the common Mediterranean β -thalassemia gene with synthetic DNA probes. *J Clin Invest* 71:775–779, 1983
18. TAN ZK, IKUTA S, HUANG T, DUGAICZYK A, ITAKURA K: Solid-phase synthesis of polynucleotides. VIII. A simplified synthesis of oligodeoxyribonucleotides. *Cold Spring Harbor Symp Quant Biol* 47:383–391, 1983
19. DAVIS RW, BOTSTEIN D, ROTH JR: Preparation of phage stocks. Purification of phage, in *A Manual for Genetic Engineering: Advanced Bacterial Genetics*, Cold Spring Harbor, N.Y., Cold Spring Harbor Laboratory, 1980, pp 74–77, 80–82
20. WALLACE RB, SCHOLD M, JOHNSON MJ, DEMBEK P, ITAKURA K: Oligonucleotide directed mutagenesis of the human β -globin gene: a general method for producing specific point mutations in cloned DNA. *Nucleic Acids Res* 9:3647–3656, 1981
21. KAZAZIAN HH JR, ORKIN SM, BOEHM CD, SEXTON JP, ANTONARAKIS SE: β -Thalassemia due to a deletion of the nucleotide which is substituted in the β^S -globin gene. *Am J Hum Genet* 35:1028–1033, 1983